6-17-2013

Numerical identification of a variable parameter in 2d elliptic boundary value problem by extragradient methods

Selin Sariaydin

Numerical Identification of a Variable Parameter in 2D Elliptic Boundary Value Problem by Extragradient Methods

by

Selin Sariaydin

A Thesis Submitted in Partial Fulfillment of the Requirements
for the Degree of Master of Science in Applied Mathematics
Department of Mathematics, College of Science

Rochester Institute of Technology
Rochester, NY

June 17, 2013

Advisor: Dr. Akhtar A. Khan

Committee: Prof. Patricia Clark

Dr. Baasansuren Jadamba

Dr. Miguel Sama

Abstract

This work focuses on the inverse problem of identifying a variable parameter in a 2-D scalar elliptic boundary value problem. It is well-known that this inverse problem is highly ill-posed and regularization is necessary for its stable solution. The inverse problem is studied in an optimization framework, which is the most suitable framework for incorporating regularization. This optimization problem is a constrained optimization problem where the constraint set is a closed and convex set of the admissible coefficients. As an objective functional, we use both the output least squares and modified output least squares functionals. It is known that the most commonly used iterative schemes for such problems require strong monotonicity of the objective functionals derivative. In the context of the considered inverse problem, this is a very stringent requirement and is achieved through a careful selection of the regularization parameter. In contrast, extragradient type methods only require the derivative of the objective functional to be monotone and this allows a greater flexibility for the selection of the regularization parameter. In this work, we use the finite element method for the discretization of the inverse problem and apply the most commonly studied extragradient methods.

Dedicated to my parents

Hava and Aziz Sariaydin
with deep love.

Acknowledgement

I would like to express my deepest gratitude to my family for supporting me throughout my entire education most recently, the completion of my masters.

I would like to thank my advisor, Dr. Akhtar Khan for his guidance. Dr. Khan provided access to an extensive set of resources including unique courses, and a chance to meet with great mathematicians from around the world.

I also consider myself very lucky to work with an excellent research group. In particular, I would like to thank Brian Winkler who was willing to spend his valuable time to overcome computational difficulties and Nate Bush who was always ready to help. Special thanks to Noorullah Maqsoodi, who was willing to study with me for hours. I owe my deepest gratitude to my good friend, Ashley Zanca, who was always there to support and encourage me to do better.

I also want to thank International Institute of Education, Turkish Fulbright Commission, and the School of Mathematical Sciences at RIT for the financial grant and the constant support.

Many thanks to my committee members: Patricia Clark, Baasansuren Jadamba, and Miguel Sama, and the director of graduate programs, Tamas Wiandt.

Contents

List of Figures

List of Tables

Chapter 1

Introduction to Inverse Problems

In the following chapter, we give a brief introduction to inverse problems. The following phrase adequately defines inverse problems:

In inverse problems one seeks unknown causes based on observation of their effects. On the other hand, for the direct problem, one seeks effects based on sufficient knowledge about the causes.

We emphasize that the inverse problems have quite different behaviour than the direct problems. Due to their special properties, most inverse problems are *ill-posed*. In 1932, J. Hadamard defined the generic properties of problems which arise from physical and natural phenomena. Based on Hadamard's definition, a mathematical problem is called *well-posed* if it has the following features:

1. **Existence:** For a suitable data set, the problem is solvable.

2. **Uniqueness:** The solution is unique.

3. **Stability:** The solution depends continuously on the data.

Following the above criteria, a problem is termed as ill-posed if it fails any of the above three conditions. However, the main concern in the study of inverse problems is the violation of the third condition, that is, the case in which the solution does not depend continuously on the data.

In this work, we focus our attention to the study of inverse problems of identifying physical parameters in 2-D elliptic partial differential equations (PDE). Inverse problems have been the subject of several papers [1, 11, 12, 13, 14, 15] .

1.1 Problem Formulation

We focus on the elliptic inverse problem of estimating the coefficient a in the elliptic boundary value problem (BVP):

$$-\nabla\cdot(a\nabla u(x)) = f \quad \text{in } \Omega, \tag{1.1a}$$

$$u = 0 \quad \text{on } \partial\Omega. \tag{1.1b}$$

This problem is a particular case of the more general following BVP

$$\nabla\cdot(a\nabla u(x,t)) = B(x)\frac{\partial u}{\partial t} + C(x,t)$$

which models a confined inhomogeneous isotropic aquifer. Here, u represents the piezometric head, a the transmissivity, $C(x,t)$ the recharge, and $B(x)$ the storativity of the aquifer. It is commonly observed that aquifers tend to be *thin* relative to their horizontal extent and thus a natural simplification is the assumption that the transmissivity varies little with depth, so that the ground water flow in these cases can be viewed as essentially two dimensional, and we can take $x = (x_1, x_2)$ in a two dimensional space. If the flow of the water has reached a steady state and we assume for simplicity that $C = 0$, then one can obtain (1.1).

There are many other models that lead to (1.1).

1.2 Variational Problem

In this work, we will use finite element methods to solve the inverse as well as the direct problem. As it is well-known, for finite element methods, the BVP should be converted into the so-called the variational form.

In this section, our objective is to introduce the variational form of (1.1). For this, we begin by recalling the product rule in multiple dimensions:

$$
\begin{aligned}
\nabla \cdot (v\nabla u) &= \nabla \cdot \left(v \begin{bmatrix} \frac{\partial u}{\partial x} \\ \frac{\partial u}{\partial y} \end{bmatrix} \right) \\
&= \nabla \cdot \left(\begin{bmatrix} v\frac{\partial u}{\partial x} \\ v\frac{\partial u}{\partial y} \end{bmatrix} \right) \\
&= \frac{\partial}{\partial x}\left(v\frac{\partial u}{\partial x} \right) + \frac{\partial}{\partial y}\left(v\frac{\partial u}{\partial y} \right) \\
&= \frac{\partial v}{\partial x}\frac{\partial u}{\partial x} + v\frac{\partial^2 u}{\partial x^2} + \frac{\partial v}{\partial y}\frac{\partial u}{\partial y} + v\frac{\partial^2 u}{\partial y^2} \\
&= \frac{\partial v}{\partial x}\frac{\partial u}{\partial x} + \frac{\partial v}{\partial y}\frac{\partial u}{\partial y} + v\left(\frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial y^2} \right) \\
&= \nabla v \cdot \nabla u + v\Delta u.
\end{aligned}
$$

Now, integrating both sides over Ω and applying the divergence theorem gives Green's identity:

$$
\begin{aligned}
\nabla \cdot (v\nabla u) &= \nabla v \cdot \nabla u + v\Delta u \\
\Rightarrow \quad \int_\Omega \nabla \cdot (v\nabla u) &= \int_\Omega \nabla v \cdot \nabla u + \int_\Omega v\Delta u \\
\Rightarrow \quad \int_{\partial\Omega} v\nabla u \cdot n &= \int_\Omega \nabla v \cdot \nabla u + \int_\Omega v\Delta u.
\end{aligned}
$$

Therefore, by denoting $\nabla u \cdot n$ with $\frac{\partial u}{\partial n}$, we get

$$
-\int_{\partial\Omega} v\Delta u = \int_\Omega \nabla v \cdot \nabla u - \int_{\partial\Omega} v\frac{\partial u}{\partial n} \tag{1.2}
$$

To write (1.1) into a variational form, we multiply it by an arbitrary $v \in H_0^1(\Omega)$ and integrate over the domain Ω. Then by using an analogue of Green's identity, we deduce the following variational form of (1.1): Find $u \in H_0^1(\Omega)$ such that

$$\int_\Omega a\nabla u \cdot \nabla v = \int_\Omega fv \text{ for all } v \in H_0^1(\Omega). \tag{1.3}$$

It is convenient to express this variational problem in terms of a trilinear form. For this, we write $V = H_0^1(\Omega)$ and $B = L^\infty(\Omega)$. Here, B is the coefficient space. We define the set of all admissible coefficients as follows:

$$A = \{a \in L^\infty(\Omega) : k_0 \leq a \leq k_1\},$$

where $k_1 > k_0 > 0$ are given.

We then define $T : B \times V \times V \to \mathbb{R}$ by

$$T(a,u,v) = \int_\Omega a\nabla u \cdot \nabla v.$$

Clearly, T is symmetric in u and v.

We further note that

$$\begin{aligned} T(a,u,v) &= \int_\Omega a\nabla u \cdot \nabla v \\ &\leq \max_{a\in\Omega} |a| \|\nabla u\|_{L^2(\Omega)} \|\nabla v\|_{L^2(\Omega)} \\ &\leq k_1 \|\nabla u\|_{L^2(\Omega)} \|\nabla v\|_{L^2(\Omega)} \end{aligned}$$

where the Cauchy-Schwarz inequality has been used. For $w \in V$, we have

$$\begin{aligned} \|\nabla w\|_{L^2(\Omega)}^2 &= \int_\Omega \nabla w \cdot \nabla w \\ &\leq \int_\Omega \left(w^2 + \nabla w \cdot \nabla w\right) \\ &= \|w\|_{H^1(\Omega)}^2 \end{aligned}$$

and consequently

$$T(a,u,v) \leq \|a\|_B \|u\|_V \|v\|_V,$$

which proves the continuity of T.

For the ellipticity of the trilinear form T, we recall the Poincaré's inequality: There exists a positive constant C, depending on the domain Ω, such that

$$\sqrt{\int_\Omega \nabla u \cdot \nabla u} \geq C\|u\|_{H^1(\Omega)} \quad \text{for all } u \in H_0^1(\Omega). \tag{1.4}$$

Consequently,

$$\begin{aligned} T(a,u,u) &= \int_\Omega a\nabla u \cdot \nabla u \\ &\geq k_0 \int_\Omega \nabla u \cdot \nabla u \\ &\geq k_0 C\|u\|_V^2 \end{aligned}$$

implying that

$$T(a,u,u) \geq \alpha\|u\|_V^2 \ \text{ for all } u \in V,\ a \in A,$$

where $\alpha = k_0 C > 0$.

This variational equation (1.3) can also be expressed as

$$T(a,u,v) = m(v) \ \text{ for all } v \in V,$$

where m is the bounded linear functional on V defined by

$$m(v) = \int_\Omega fv.$$

The above arguments can easily be extended to the BVP with mixed boundary conditions such as:

$$\begin{aligned} -\nabla \cdot (a\nabla u) &= f \ \text{ in } \Omega, \\ u &= 0 \ \text{ on } \Gamma_1, \\ a\nabla u(x) \cdot \nu &= g \ \text{ on } \Gamma_2 \end{aligned}$$

where $\partial\Omega = \Gamma_1 \cup \Gamma_2$.

For this case, the underlying Hilbert space becomes:

$$V = \{v \in H^1(\Omega) : \quad v = 0 \text{ on } \Gamma_1\}.$$

The variational form, by using standard arguments, becomes:

$$\int_\Omega a\nabla u\cdot\nabla v = \int_\Omega fv + \int_{\Gamma_2} vg \quad \text{for all } v\in V.$$

Therefore, in an attempt to write it in the form (1.3), the trilinear form remains the same as for (1.1) but the functional m has to be modified to:

$$m(v) = \int_\Omega fv + \int_{\Gamma_2} vg.$$

1.3 Solution Differentiability

The existence and uniqueness of solution to the variational form of the problem can be given by the Lax-Milgram theorem that can be found in the appendix. Therefore, we can define $\mathscr{F} : A \to V$ such that $\mathscr{F}(a) = u$ is the solution of the variational problem.

Lemma 1.3.1. *For each $a\in A$, $u=\mathscr{F}(a)$ satisfies*

$$\|u\|_V \le \alpha^{-1}\|m\|_{V^*}.$$

Proof. Using the V-ellipticity of the trilinear form:

$$\begin{aligned}
& T(a,u,v) = m(v) \ \text{ for all } v\in V \\
\Rightarrow\ & T(a,u,u) = m(u) \\
\Rightarrow\ & \alpha\|u\|_V^2 \le \|m\|_{V^*}\|u\|_V \\
\Rightarrow\ & \|u\|_V \le \alpha^{-1}\|m\|_{V^*}.
\end{aligned}$$

□

Theorem 1.3.2. *The solution operator $\mathscr{F}$ satisfies the following conditions:*

$$\begin{aligned}
\|\mathscr{F}(a)-\mathscr{F}(b)\|_V &\le \frac{\beta}{\alpha}\|\mathscr{F}(a)\|_V\|b-a\|_B, \\
\|\mathscr{F}(a)-\mathscr{F}(b)\|_V &\le \frac{\beta}{\alpha}\|\mathscr{F}(b)\|_V\|b-a\|_B, \\
\|\mathscr{F}(a)-\mathscr{F}(b)\|_V &\le \frac{\beta}{\alpha^2}\|m\|_{V^*}\|b-a\|_B.
\end{aligned}$$

The following result gives information regarding the differentiability of the solution map:

Theorem 1.3.3. *For each a in the interior of A, the operator $\mathscr{F}$ is differentiable at a, and $\delta u = D\mathscr{F}(a)\delta a$ is the unique solution to the variational problem.*

$$T(a, \delta u, v) = -T(\delta a, u, v) \ \textit{for all } v \in V, \tag{1.5}$$

where $u = \mathscr{F}(a)$. Moreover,

$$\|D\mathscr{F}(a)\| \leq \frac{\beta}{\alpha}\|u\|_V,$$

and hence, for all a in the interior of Y,

$$\|D\mathscr{F}(a)\| \leq \frac{\beta}{\alpha^2}\|m\|_{V^*}.$$

Now, we can also show that $\mathscr{F}$ is infinitely differentiable. We will just state the second derivative.

Theorem 1.3.4. *For each a in the interior of A, the operator $\mathscr{F}$ is twice-differentiable at a, and*

$$\delta^2 u = D^2\mathscr{F}(a)(\delta a_1, \delta a_2)$$

is the unique solution to the variational problem.

$$T(a, \delta^2 u, v) = -T(\delta a_2, D\mathscr{F}(a)\delta a_1, v) - T(\delta a_1, DF(a)\delta a_2, v) \ \textit{for all } v \in V. \tag{1.6}$$

Moreover,

$$\left\|D^2\mathscr{F}(a)\right\| \leq \frac{2\beta^2}{\kappa^2}\|\mathscr{F}(a)\|_V \leq \frac{2\beta^2}{\kappa^3}\|m\|_{V^*}.$$

Further details can be found in [12].

1.4 An Optimization Framework of the Inverse Problem

We recall that there are two primary approaches for solving the inverse problem of identifying the coefficient, a. The first approach reformulates the inverse problem as an optimization problem and then employs some suitable method for the solution. The second approach views (1.1) as a hyperbolic partial differential equation in a.

In this work, we pose the considered inverse problem as an optimization problem whose numerical solution is an approximation of the coefficient to be identified. As is well-known, the inverse problem is ill-posed, and some type of regularization is necessary. Since this is more easily accomplished in the optimization setting, this class of methods has been the subject of most of the research.

In the following, we briefly outline some of the main approaches available in the literature for the inverse problems.

1.4.1 Output Least-Squares

A common approach for solving parameter identification problems is the output least-squares (OLS) approach. Applied to the elliptic inverse problem of finding a in (1.1), the OLS approach minimizes the functional

$$J_1(a) = \|u(a) - z\|^2, \tag{1.7}$$

where z is the data (the measurement of u), $\|\cdot\|$ is a suitable norm and u solves the BVP or its variational form: Find $u \in H_0^1(\Omega)$ such that

$$\int_\Omega a \nabla u \nabla v = \int_\Omega f v \quad \text{for all} \quad v \in H_0^1(\Omega). \tag{1.8}$$

1.4.2 Modified Output Least-Squares

For the scalar inverse problem, a variation on the OLS approach was proposed independently by Knowles [22]. They replaced the L_2 norm by the coefficient-dependent energy norm:

$$J_2(a) = \frac{1}{2}\int_\Omega a\nabla\left(u(a)-z\right)\cdot\nabla\left(u(a)-z\right). \tag{1.9}$$

In [12], the following extension of the modified output least-squares (MOLS) functional was introduced:

$$J_2(a) = \frac{1}{2}T(a, u-z, u-z),$$

where T is a general trilinear form.

1.4.3 Regularization

Regularization is a very common technique to solve ill-posed problems or prevent overfitting. Regularization has an impact role in uniqueness. It can make a non-unique problem become a unique problem. Moreover, it is important to choose regularization parameter correctly since choosing a poor regularization parameter can create a different problem. We can choose one of the following forms for the regularization:

$$\begin{aligned} L_2\ norm: \quad R_\varepsilon(a) &= \frac{\varepsilon}{2}||a||^2_{L_2} \\ H_1\ semi-norm: \quad R_\varepsilon(a) &= \frac{\varepsilon}{2}||\nabla a||^2_{L_2} \\ H_1\ norm: \quad R_\varepsilon(a) &= \frac{\varepsilon}{2}(||a||^2_{L_2} + ||\nabla a||^2_{L_2}) \end{aligned} \tag{1.10}$$

where $R_\varepsilon(a)$ is the regularization.

The basic idea is to add a suitable functional to penalize numerical features that are not natural to the variational problem. Moreover, regularization has significant importance on smoothing the data, hence it is very critical to choose the right regularization parameter. In the Performance Analysis chapter, we include some figures that show how the computed solution can differ depending on the regularization parameter.

Now, we can define the constrained optimization problem as following:

$$\min_{a \in A}\{J(a) + R_{\varepsilon}(a)\} \tag{1.11}$$

where ε is the regularization parameter and $J(a)$ is either OLS, $J_1(a)$, or MOLS, $J_2(a)$.

We will solve the inverse problem by writing the minimization problem as a variational inequality of finding $a^* \in A$ such that

$$\langle DJ(a^*) + DR_{\varepsilon}(a^*), a - a^* \rangle \geq 0 \ \ \forall a \in A,$$

Here, OLS is non-convex and the following relaxed-monotonicity holds [21]:

$$\langle DJ_1(x) - DJ_1(y), x - y \rangle \geq -m||x-y||^2, \ \ \forall x, y \in A, \, m > 0$$

For a sufficiently smooth and strongly convex regularization term R_{ε}, we have

$$\langle DR_{\varepsilon}(x) - DR_{\varepsilon}(y), x - y \rangle \geq \varepsilon m_0 ||x-y||^2, \ \ \forall x, y \in A, \, m_0 > 0$$

To use extragradient methods which are convergent under strong monotonicity, we need to assume that

$$-m + \varepsilon m_0 = m_1 > 0$$

While $m_0 = 1$ and m is usually fixed, regularization parameter, ε, is required to be large for convergence [26]. It is also useful to state that if ε is too large, then we will have an over-regularized solution and choosing a small ε will lead to an under-regularized solution.

Chapter 2

The Finite Element Method for the Inverse Problem

In this chapter, we will give the basic idea of the finite element method (FEM). To obtain the finite element method for the two dimension problem, we will use the weak form of the boundary value problem. Then, we will apply the Galerkin method to the weak form of the problem. Finally, we will define the finite element spaces which is the most important difference from the one dimension case. The details can be found in [9].

Recall the weak form of the 2D problem:

$$\int_\Omega a(x)\nabla u\cdot\nabla v = \int_\Omega fv \;\; for\ all\ v\in V. \tag{2.1}$$

Now we can define the bilinear form of $a(\cdot,\cdot)$ as the following:

$$a(u,v) = (f,v)$$

where

$$a(u,v) = \int_\Omega a(x)\nabla u\cdot\nabla v$$

and

$$(f,v) = \int_\Omega fv$$

Hence, the problem is to find $u \in V$ such that

$$a(u,v) = (f,v) \quad \forall v \in V \tag{2.2}$$

2.1 The Galerkin Method

The Galerkin method approximates to the solution accurately when energy inner product is used and the variational form of the BVP computes the value of $a(u,v)$ for all $v \in V$. Solving $KU = F$ matrix-vector equation gives the best approximation to u where K is the stiffness matrix and F is the load vector.

Let V_n be finite dimensional subspace of V and let $\{\phi_1, \phi_2, \cdots, \phi_n\}$ be a basis for V_n. Then (2.2) is defined as the Galerkin problem [9] :

$$u_n = \sum_{j=1}^{n} U_j \phi_j \tag{2.3}$$

Then, we can rewrite the Galerkin problem as

$$a(u_n, \phi_i) = (f, \phi_i) \quad for\ i = 1,2,3,...,n \tag{2.4}$$

Using (2.3), we have

$$\begin{aligned} & a(\sum_{j=1}^{n} U_j \phi_j, \phi_i) = (f, \phi_i) \\ & \Rightarrow \sum_{j=1}^{n} a(\phi_j, \phi_i) U_j = (f, \phi_i) \ \ for\ \ i = 1,2,3,...,n \\ & \Rightarrow KU = F \end{aligned} \tag{2.5}$$

where K is the stiffness matrix and f is the load vector, that is:

$$K_{ij} = a(\phi_j, \phi_i) \ \ and \ \ f_i = (f, \phi_i) \ \ i, j = 1,2,3,...,n$$

Then, the vector U defines the approximate solution.

FEM can be seen as a Galerkin method when a particular subspace and its basis are chosen. If the coefficient matrix is dense, solving the linear system and finding the orthogonal basis are difficult since the computations require extensive time. However, FEM uses a basis that leads to a sparse coefficient matrix. Since the sparse matrix has mostly zero entries, solving the linear system will be less time-consuming.

Another interpretation of the Galerkin method can be defined as in an abstract setting the variational problem of finding $u \in V$ such that

$$a(u,v) = f(v) \quad \forall v \in V$$

is given by

$$J(u) = \frac{1}{2}a(u,u) - f(u).$$

The problem is find $u \in V$ such that

$$\min_{u\in V} J(u)$$

Instead we solve the problem

$$\min_{w\in W} J(w)$$

where W is a finite dimensional subspace of V. This approach is called the Ritz Method.

Let $\{w_1, \cdots, w_n\}$ be a basis for W. Then, we have

$$w = \sum_{i=1}^{n} U_i w_i$$

Therefore,

$$\begin{aligned}
J(w) &= \frac{1}{2}a(w,w) - f(w) \\
&= \frac{1}{2}a(\sum_{j=1}^{n} U_j w_j, \sum_{i=1}^{n} U_i w_i) - f(\sum_{i=1}^{n} U_i w_i) \\
&= \frac{1}{2}\sum_{i=1}^{n}\sum_{j=1}^{n} a(w_j, w_i)U_j U_i - \sum_{i=1}^{n} f(w_i)U_i \\
&= \frac{1}{2}U \cdot KU - F \cdot U
\end{aligned}$$

where K is the stiffness matrix and F is the load vector.

Therefore, the problem is to find $U \in R^n$ such that

$$J(U) = \frac{1}{2} < U, KU > - < F, U >$$

is minimized.

The gradient of J can be defined as

$$\nabla J(U) = KU - F$$

and $\nabla J(U) = 0$ then the linear system is given by

$$KU = F$$

Recall the stiffness matrix, K has the following form:

$$K_{ij} = a(\phi_j, \phi_i) = \sum_{i=1}^{n} a(\phi_j, \phi_i) \tag{2.6}$$

Then, the stiffness matrix can be written as

$$K_{ij} = \int_{\Omega} k \nabla \phi_i \cdot \phi_j \;\; i, j = 1, 2, \cdots, n$$

[10] includes details about the theory and some examples to show how to compute the stiffness matrix, K.

The load vector has the following form:

$$\begin{aligned}
F_i &= \langle f, \phi_i \rangle \\
&= \int_{\Omega} f \phi_i, \quad i = 1, 2, \cdots, n
\end{aligned}$$

2.2 Discrete Formulas for the OLS

We will need discrete analogue of the OLS functional to solve the inverse problem numerically, so we will define the finite element solution operator as $\mathscr{F}: \mathbb{R}^m \to \mathbb{R}^n$ mapping a coefficient $a \in A_m$ to the approximate solution $u \in U_m$ where $m = n+2$. Then $\mathscr{F}(A) = U$ where U is defined as

$$K(A)U = F$$

$K(A)$, the stiffness matrix, has the following form:

$$K_{ij} = \int_\Omega k\nabla\phi_j \cdot \nabla\phi_i \;\; i = 1,2,\cdots,n$$

and the load vector has the following form:

$$F_{ij} = \int_\Omega f\phi_i \;\; i = 1,2,\cdots,n$$

It is also important to know that the tensor is defined as

$$T_{ijk} = \int_\Omega \psi_k\nabla\phi_i \cdot \nabla\phi_j \;\; i,j = 1,2,\cdots,n, \;\; k = 1,2,\cdots,m$$

Therefore, the stiffness matrix can be defined by

$$K(A)_{ij} = T_{ijk}A_k$$

Now, we can discretize the system using the basis functions for A_n, $\{\psi_1,\psi_2,\cdots,\psi_m\}$ and for U_n, $\{\phi_1,\phi_2,\cdots,\phi_n\}$. Thus,

$$a = \sum_{i=1}^{m} A_i\psi_i$$

$$u = \sum_{i=1}^{n} U_i\phi_i$$

Therefore, the discrete form of the OLS can be stated as [15]

$$J_1(A) = \frac{1}{2}(U-Z)\cdot M(U-Z)$$

where the mass matrix $M \in \mathbb{R}^{n\times n}$ is

$$M_{i,j} = \int_\Omega \phi_i \phi_j dx, \quad i,j = 1,2,\cdots,n$$

We will also need the gradient of the objective functional, that is:

$$\nabla J_1(A) = -L(U)^T K(A)^{-1} M(U-Z)$$

where the the adjoint-stiffness matrix L satisfies the following condition:

$$L(U)A = K(A)U, \ \forall A \in \mathbb{R}^n, \ U \in \mathbb{R}^n$$

We can also compute the hessian matrix by using similar notation:

$$\begin{aligned}\nabla^2 J_1(A) = {} & L(U)^T K(A)^{-1} M K(A)^{-1} L(U) + L(K(A))^{-1} M(U-Z)^T K(A)^{-1} L(U) \\ & + L(U)^T K(A)^{-1} L(K(A))^{-1} M(U-Z))\end{aligned}$$

2.3 Discrete Formulas for the MOLS

Define an objective function $J_2 : \mathbb{R}^m \to \mathbb{R}$ as

$$J_2(A) = \frac{1}{2}\int_\Omega a(\nabla u - \nabla z)\cdot(\nabla u - \nabla z)$$

where u is defined as $\mathscr{F}(A) = U$ and z is the measurement of the exact solution u of the original problem.

The discrete modified output least squares can be written as

$$J_2(A) = \frac{1}{2}(U-Z)\cdot K(A)(U-Z)$$

and it follows that

$$\nabla J_2(A) = -\frac{1}{2}L(U)^T U + \frac{1}{2}L(Z)^T Z$$

and the hessian can be defined as

$$\nabla^2 J_2(A) = L(U)^T K(A)^{-1} L(U)$$

The details about the derivation of the discrete form for the OLS and MOLS can be found in [15].

2.4 Discrete Formulas for the Regularization

The discrete formulas for the regularization, $R_\varepsilon(a)$ are as the following:

$$
\begin{aligned}
L_2\ norm: \quad & R_\varepsilon \alpha = \frac{\varepsilon}{2} A^T \tilde{M} A \\
H_1\ semi-norm: \quad & R_\varepsilon \alpha = \frac{\varepsilon}{2} A^T \tilde{K} A \\
H1\ norm: \quad & R_\varepsilon \alpha = \frac{\varepsilon}{2} A^T (\tilde{M} + \tilde{K}) A
\end{aligned}
$$

where $\tilde{M}$, and $\tilde{K}$ represent mass and stiffness matrix of size $(n+2)^2 \times (n+2)^2$, respectively.

Chapter 3

Extragradient Methods

3.1 Literature Review

In 1976, the first extragradient method was introduced by Korpelovich [23]. To describe the framework of her contribution, let f be a convex-concave function with a non-empty set of saddle points on PQ, where P and Q are convex closed subsets of finite-dimensional space. She assumed that the function f is differentiable and f' satisfied the Lipschitz condition with some constant L. Her work was motivated by the fact that the gradient method with constant stepsize under those conditions is in general known not to converge to the set of saddle points. As a remedy, she proposed a scheme defined by recursive relations by projecting twice on the underlying convex sets and proved the convergence of the generated sequence to some saddle point. She showed that in a particular case, when f is the Lagrange function of linear programming problem, the rate of convergence is exponential.

In 1987, Khobotov [18] presented a modification of the extragradient method proposed by Korpelovich [23] for solving variational inequalities defined for continuous monotone operators in finite dimensional spaces. He proved the convergence of the proposed scheme and discussed interesting applications in areas such as convex optimization, min-max theory and game theory.

Marcotte [24] strengthened the method proposed by Khobotov by providing some useful strategies for its implementation.

In 1996, Solodov and Tseng introduced a modification to projection-type methods by using a strongly monotone operator.

In addition, a nice geometric interpretation of the extragradient methods was proposed in [16, 17]. The method is continuous and satisfies a certain generalized monotonicity assumption (e.g., it can be pseudomonotone). Later, this method was developed by Solodov and Svaiter [29]. The idea of the method is to construct an appropriate hyperplane which strictly separates the current iterate from the solutions of the problem. This procedure requires a single projection onto the feasible set and employs an Armijo-type linesearch along a feasible direction. Then the next iterate is obtained as the projection of the current iterate onto the intersection of the feasible set with the halfspace containing the solution set. The method is globally convergent to a solution of the variational inequality problem.

He [30] implemented an extension of the Goldstein's projection method which was later improved in [31]. This method provided an easily-implementable Armijo-type strategy using the scaling parameters.

Popov [27] proposed a regularized extragradient method for solving a variational inequality with monotone Lipschitz operator. The inequality was considered in a finite-dimensional Euclidean space. The method allowed to construct an iterative process that converged to a solution of minimal norm. The peculiarity of the proposed process is that the descent direction at every iteration step is calculated only once (not twice as in the standard extragradient method). In the case when the inequality is unsolvable, it was shown that the proposed method generated some perturbed (already solvable) problem. The latter is the ”best approximation" (in the sense suggested by the author) to the given problem.

Recently, Y. Censor, A. Gibali, S. Reich [8] proposed two extensions of the well-known extragradient method for variational inequality problems. The first extended method replaced the second orthogonal projection in the original extragradient method by a specific subgradient projection and the second extended method allowed projections onto the members of an infinite sequence of subsets which epi-converges to the feasible set of the VIP. Both methods were shown to be convergent under suitable conditions.

In [7], the authors presented a subgradient extragradient method for solving variational inequalities in Hilbert space and proved the weak-convergence of the method.

Extragradient methods have been well-studied in various papers.(see [2, 3, 4, 5, 25, 33, 34, 35])

3.2 The Projection Method

Let $J : A \to H$ be a continuous function and consider the variational inequality problem of finding x^* such that

$$x^* \in A \quad \langle J'(x^*), x - x^* \rangle \geq 0, \forall x \in A \tag{3.1}$$

where A is a nonempty convex set. One of the most common methods is to solve the variational inequality problem by using the projection algorithm:

$$x^{k+1} = P_A(x^k - \alpha J'(x^k)) \tag{3.2}$$

where $P_A(\cdot)$ is the orthogonal projection map onto A, and α is the step length. Here, projection method uses the gradient of the objective function as the direction. x^* is a solution to (3.1) if and only if

$$x^* = P_A(x^* - \alpha J'(x^*)) \tag{3.3}$$

Lemma 3.2.1. *Let A be a closed convex subset of Hilbert space. Let J' be Lipschitz continuous and strongly monotone in A, that is:*

$$\begin{aligned} ||J'(x) - J'(y)|| &\leq L||x-y|| \ \ \forall x,y \in A && (3.4)\\ \langle J'(x) - J'(y), x-y\rangle &\geq l\,||x-y||^2 \ \ \forall x,y \in A && (3.5)\end{aligned}$$

Then, the projection method (3.2) *converges to a solution of* (3.1) *if* $\alpha \in (0, 2l/L^2)$ *where* l *is the monotonicity constant and* L *is the Lipschitz constant.*

Proof. We can prove the convergence of the projection method using contractive properties of the operator. Set $x \to x^{k+1} - \alpha J'(x^{k+1})$ and $y \to x^* - \alpha J'(x^*)$ in $||P_A(x) - P_A(y)|| \leq ||x-y||$ for all $x, y \in H$ then:

$$\begin{aligned} \|x^{k+1} - x^*\|^2 &= \|P_A(x^k - \alpha J'(x^k) - P_A(x^* - \alpha J'(x^*)\|^2 \\ &\leq \|(x^k - \alpha J'(x^k)) - (x^* - \alpha J'(x^*))\|^2 \\ &= \|x^k - x^*\|^2 - 2\alpha\langle J'(x^k) - J'(x^*), x^k - x^*\rangle + \alpha^2\|J'(x^k) - J'(x^*)\|^2 \\ &\leq (1 - 2\alpha l + \alpha^2 L^2)\|x^k - x^*\|^2, \end{aligned}$$

which due to the assumption that $\alpha < \frac{2l}{L^2}$ ensures that

$$(1 - 2\alpha l + \alpha^2 L^2) < 1$$

and consequently, we get that the sequence $\{x^k\}$ converges strongly to x^*. □

Projection methods are easy to implement. However, there are many drawbacks of the iterative algorithm. First and foremost, J' is required to be strongly monotone and Lipschitz continuous which are very powerful criteria to meet. Moreover, to find an appropriate step length might be difficult since we do not know the values of l, L and failure to choose these constants may lead slow convergence or not convergence at all.

A mapping is coercive with $\grave{c} = 1/L^2$ if it is strongly monotone and Lipschitz continuous and any coercive mapping is monotone and Lipschitz with $L = 1/\grave{c}$ such that

$$\langle J'(x) - J'(y), x - y \rangle \geq 0, \quad \forall \ x, y \in A.$$

The projection algorithm converges for coercive variational inequalities which relaxes the requirement on J' [32]. However, the problem of choosing the right initial step length still exists since it still depends on the Lipschitz constant.

One modification can be to estimate Lipschitz constant by introducing adaptivity in the projection method by Armijo line search. This modification allows step length to vary at each iteration and relaxes the requirements on J'. Another modification can be projecting twice at each iteration which is called *extragradient methods* that is first introduced by Korpelevich [23].

3.3 Extragradient Methods

The focus of this research is to implement the extragradient methods for the elliptic boundary value problem which relaxes the convergence requirements. In this section, we will give brief introduction and convergence analysis for some extragradient methods. Then, we will discuss computational aspects of the each method and identify a variable parameter in 2D elliptic boundary value problem by these methods.

The projection methods requires strong theoretical properties for convergence that are not easy to verify in practice. Korpelevich [23] proposes a modified version to the projection method to relax these requirements. Let us recall the projection method as following:

$$\bar{x}^k = P_A(x^k - \alpha J'(x^k)) \tag{3.6}$$

Here, the gradient of the point $\bar{x}^k$ will be the direction for the new point. Overall, the basic idea is to project twice at each iteration to find the solution of (3.1) for given $x^0 \in A$ such that

$$\begin{aligned} \bar{x}^k &= P_A(x^k - \alpha J'(x^k)) && \text{(3.7a)} \\ x^{k+1} &= P_A(x^k - \alpha J'(\bar{x}^k)), && \text{(3.7b)} \end{aligned}$$

where α is the constant step length.

Theorem 3.3.1 (Korpelevich, [23])**.** *Let A be a closed convex subset of $\mathbb{R}^n$ and A^* be a nonempty set of solutions of* (3.1)*. J' is monotone operator in A and Lipschitz continuous with $\alpha \in (0, 1/L)$ where L is the Lipschitz constant, then x^k that defined by* (3.7) *converges to some solution of the variational inequality, $x^* \in A^*$.*

Choosing the right step length plays a vital role for convergence as it is stated in the Theorem 3.3.1. Since L is unknown, it is difficult to determine a suitable step length, α. Placing a small value for step length may lead slow convergence. However, the method might not converge if the step length is too large.

Then, Khobotov, [18] proposed a modification to choose step length which is changing α at each iteration:

$$\begin{aligned} \bar{x}^k &= P_A(x^k - \alpha_k J'(x^k)) && \text{(3.8a)} \\ x^{k+1} &= P_A(x^k - \alpha_k J'(\bar{x}^k)). && \text{(3.8b)} \end{aligned}$$

If $J'(\bar{x}^k) = 0$, then $\bar{x}^k$ is a solution to the variational problem of (3.1) and here α_k satisfies the following inequality:

$$0 < \alpha_k \leq \min\left\{ \bar{\alpha}, \varepsilon \frac{||x^k - \bar{x}^k||}{||J'(x^k) - J'(\bar{x}^k||} \right\} \tag{3.9}$$

where $\bar{\alpha}$ is the maximum value for α and $\varepsilon \in (0,1)$. Khobotov states the following theorem which is a very important for convergence analysis of the extragradient method.

Theorem 3.3.2 (Khobotov, [18])**.** *Let A be a closed convex subset of $\mathbb{R}^n$ and A^* be a nonempty set of solutions of* (3.1)*.where J' is a continuous monotonic operator in A. $x^0 \in A$ and x_k is the sequence that obtained by* (3.7)*. Then, the following inequality holds for any nonnegative x_k and $x^* \in A^*$:*

$$||x^{k+1} - x^*||^2 \leq ||x^k - x^*||^2 - ||x^k - \bar{x}^k||^2 \left\{ 1 - \alpha_k^2 \frac{||J'(x^k) - J'(\bar{x}^k)||^2}{||x^k - \bar{x}^k||^2} \right\} \tag{3.10}$$

Proof. For any u and $v \in A$ vectors,

$$\langle u - P_A(u), v - P_A(u) \rangle \leq 0 \tag{3.11}$$

Thus,

$$\begin{aligned} ||u - v||^2 &= ||u - P_A(u) + P_A(u) - v||^2 \\ &= ||u - P_A(u)||^2 - 2\langle u - P_A(u), v - P_A(u) \rangle + ||v - P_A(u)||^2 \\ &\geq ||u - P_A(u)||^2 + ||v - P_A(u)||^2. \end{aligned} \tag{3.12}$$

By setting $u = x^k - \alpha_k J'(\bar{x})^k$, $v = x^*(x^* \in A^*)$, $P_A(u) = x^{k+1}$, we have

$$\begin{aligned} ||x^k - \alpha_k J'(\bar{x}^k) - x^*||^2 &\geq ||x^k - \alpha_k J'(\bar{x})^k - x^{k+1}||^2 + ||x^* - x^{k+1}||^2 \qquad (3.13) \\ ||x^{k+1} - x^k||^2 &\leq ||x^k - \alpha_k J'(\bar{x}^k) - x^*||^2 - ||x^k - \alpha_k J'(\bar{x})^k - x^{k+1}||^2 \\ ||x^{k+1} - x^k||^2 &\leq ||x^k - x^*||^2 - 2\alpha_k \langle x^k - x^*, J'(\bar{x}^k) \rangle + ||\alpha_k J'(\bar{x}^k)||^2 - \\ &+ ||x^k - x^{k+1}||^2 + 2\alpha_k \langle x^k - x^{k+1}, J'(\bar{x}^k) \rangle - ||\alpha_k J'(\bar{x}^k)||^2 \\ ||x^{k+1} - x^k||^2 &\leq ||x^k - x^*||^2 - ||x^k - x^{k+1}||^2 + 2\alpha_k \langle x^* - x^{k+1}, J'(\bar{x}^k) \rangle. \end{aligned}$$

The following inequality is trivial:

$$\begin{aligned} \langle x^* - x^{k+1}, J'(\bar{x}^k) \rangle &= \langle x^* - \bar{x}^k, J'(\bar{x}^k) \rangle + < \bar{x}^k - x^{k+1}, J'(\bar{x}^k) > \\ &\leq \langle \bar{x}^k - x^{k+1}, J'(\bar{x}^k) \rangle \end{aligned} \tag{3.14}$$

J' monotone and $x^* \in A^*$. Therefore, plug (3.14) into (3.13):

$$\begin{aligned}
||x^{k+1}-x^k||^2 &\leq ||x^k-x^*||^2 - ||x^k-\bar{x}^k+\bar{x}^k-x^{k+1}||^2 + 2\alpha_k\langle \bar{x}^k-x^{k+1}, J'(\bar{x}^k)\rangle \\
&= ||x^k-x^*||^2 - ||x^k-\bar{x}^k||^2 - ||\bar{x}^k-x^{k+1}||^2 - 2\langle x^k-\bar{x}^k, \bar{x}^k-x^{k+1}\rangle \\
&+ 2\alpha_k\langle \bar{x}^k-x^{k+1}, J'(\bar{x}^k)\rangle \\
&= ||x^k-x^*||^2 - ||x^k-\bar{x}^k||^2 - ||\bar{x}^k-x^{k+1}||^2 + 2\langle x^k-\alpha_k J'(\bar{x}^k)-\bar{x}^k, x^{k+1}-\bar{x}^k\rangle \\
&= ||x^k-x^*||^2 - ||x^k-\bar{x}^k||^2 - ||\bar{x}^k-x^{k+1}||^2 + 2\langle x^k-\alpha_k J'(x^k)-\bar{x}^k, x^{k+1}-\bar{x}^k\rangle \\
&+ 2\alpha_k\langle J'(x^k)-J'(\bar{x}^k), x^{k+1}-\bar{x}^k\rangle \\
&\leq ||x^k-x^*||^2 - ||x^k-\bar{x}^k||^2 - ||\bar{x}^k-x^{k+1}||^2 + 2\alpha_k||J'(x^k)-J'(\bar{x}^k)||\,||x^{k+1}-\bar{x}^k||
\end{aligned} \tag{3.15}$$

Setting $v = x^{k+1}$ and $u = x^k - \alpha_k J'(x^k)$ and $P_A(u) = \bar{x}^k$ in (3.11):

$$\langle x^k-\alpha_k J'(x^k)-\bar{x}^k, x^{k+1}-\bar{x}^k\rangle \leq 0 \tag{3.16}$$

For any $x^{k+1}, x^k, \bar{x}^k, \alpha_k$, we have

$$||x^{k+1}-x^*||^2 + {\alpha_k}^2||J'(x^k)-J'(\bar{x}^k)||^2 \geq 2\alpha_k||J'(x^k)-J'(\bar{x}^k)||\,||x^{k+1}-\bar{x}^k|| \tag{3.17}$$

then, it follows

$$\begin{aligned}
||x^{k+1}-x^*||^2 &\leq ||x^k-x^*||^2 - ||x^k-\bar{x}^k||^2 - ||\bar{x}^k-x^{k+1}||^2 + ||x^{k+1}-\bar{x}^k||^2 + {\alpha_k}^2||J'(x^k)-J'(\bar{x}^k)||^2 \\
&\leq ||x^k-x^*||^2 - ||x^k-\bar{x}^k||^2\left\{1-\alpha_k^2\frac{||J'(x^k)-J'(\bar{x}^k)||^2}{||x^k-\bar{x}^k||^2}\right\}
\end{aligned} \tag{3.18}$$

Further details can be found in [18]]. □

Any limiting point x^* satisfies $x^* = P_A(x^* - \alpha_k J'(x^*))$ since the projection operator is continuous. x^* must be a solution to the variational problem when $\{\alpha_k\}$ stays away from zero.

Khobotov also points out that extragradient methods are effective in finding the solutions of variational inequality problems even if the operator J' is not strictly monotonic. In addition, the iterative algorithm for step length removes the requirement of Lipschitz continuity of J'.

Moreover, Marcotte [24] introduces a new approach to reduce step length. Recall the inequality:

$$||x^{k+1} - x^*||^2 \leq ||x^k - x^*||^2 - ||x^k - \bar{x}^k||^2 \left\{ 1 - \alpha_k^2 \frac{||J'(x^k) - J'(\bar{x}^k)||^2}{||x^k - \bar{x}^k||^2} \right\} \tag{3.19}$$

Here, the convergence analysis does not depend on the contraction argument. However, we should still estimate α such that x^* will be a solution to the variational problem. The idea is to minimize the right hand side of (3.19), thus we can choose step length such that

$$\left\{ 1 - \alpha_k^2 \frac{||J'(x^k) - J'(\bar{x}^k)||^2}{||x^k - \bar{x}^k||^2} \right\}$$

is maximized. The step length is given as

$$\alpha_k = \frac{1}{\sqrt{2}} \frac{||x^k - \bar{x}^k||}{||J'(x^k) - J'(\bar{x}^k)||}$$

See [24] for more details.

Another extragradient method that we will study is one of the projection-contraction methods that is introduced by Solodov and Tseng [28]. A scaling matrix, M symmetric and positive definite, is introduced to accelerate the convergence. These methods only require one projection at each iteration and define a general operator at the second projection such that

$$x^{k+1} = x^k - \gamma M^{-1}(T_\alpha(x^k) - T_\alpha P_A(x^k - \alpha_k J'(x^k))) \tag{3.20}$$

where $\gamma > 0$ and $T_\alpha = (I - \alpha J')$. Here, the operator T_α is strongly monotone and α is chosen by Armijo line search. The method is convergent when the solution set A^* is nonempty and J' is monotone [28].

Furthermore, Goldstein's extragradient methods [30, 31] have been widely used to solve variational inequality problems. In particular, we will implement an improved version of Goldstein's method for our problem [31] that chooses the step size along the descent direction $r(x^k, \beta_k)$. The scaled residue of the projection method is defined in the following:

$$r(x^k, \beta_k) = \frac{1}{\beta_k}\{J'(x^k) - P_A[J'(x^k) - \beta_k x^k]\} \tag{3.21}$$

If $x^* \in A^*$ then $r(x^*, \beta) = 0$.

Theorem 3.3.3. *Let A be a closed convex subset of $\mathbb{R}^n$ and A^* be a nonempty set of solutions of* (3.1)*. Let x^* be an arbitrary point in A^*, $\gamma \in (0,2)$ and $\beta_k \in [\beta_L, \beta_U] \subset (1/(4\tau), +\infty)$. For given x^k, the new iterate is*

$$x^{k+1} = x^k - \gamma\,\alpha_k\, r(x^k, \beta_k)$$

then we have

$$||x^{k+1} - x^*|| \leq ||x^k - x^*||^2 - \frac{2-\gamma}{\gamma}||x^k - x^{k+1}||^2,$$

and

$$||x^k - x^{k+1}|| \leq \gamma||r(x^k, \beta_k)||$$

This theorem shows x^{k+1} closer to a solution at each iteration for any $\gamma \in (0,2)$. Thus, it converges to a solution of the variational inequality problem. Proof and further convergence analysis of the method can be found in [31].

The final extragradient method that we implement is Hyperplane method [16, 32] of the following form:

$$\bar{x}^k = P_A(x^k - \alpha_k J'(x^k)) \qquad x^{k+1} = P_A(x^k - \lambda_k J'(\bar{x}^k)) \tag{3.22}$$

where $\alpha_k > 0$ is located through a finite bracketing procedure and

$$\lambda_k = \frac{\langle J'(\bar{x}^k), \bar{x}^k - x^k\rangle}{||J'(\bar{x}^k)||^2}.$$

Let $H_k = \{x \in \mathbb{R}^n : \langle J'(\bar{x}^k), \bar{x}^k - x^k\rangle = 0\}$ be an hyperplane normal to $J'(\bar{x}^k)$.

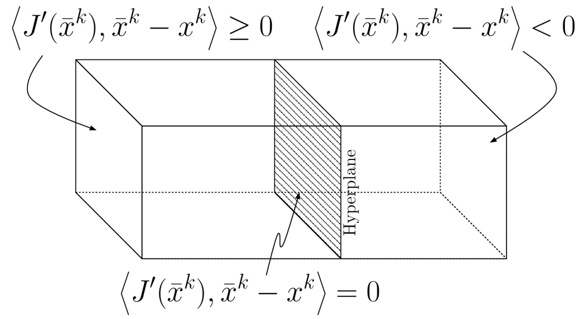


Figure 3.1: Geometry of Hyperplane Extragradient Method

As it shows in the Figure 3.1, one side includes all the solutions of the variational problem. A hyperplane excludes x^k that satisfies the following condition:

$$\langle J'(\bar{x}^k), \bar{x}^k - x^k \rangle < 0$$

It can be concluded that $||x^{k+1} - x^*|| \leq ||x^k - x^*||$ for any solution x^* of the variational problem. Let J' be monotone and continuous, and the solution set A^* be nonempty, then the hyperplane method converges [16].

Now, we can move forward to computational aspects of the extragradient methods that we discussed above. We will state the step algorithms for each method and implement for the following test problem:

Test Problem

$$\begin{aligned} a &= 1 + x(x-1) + y(y-1) \\ u &= x(e^x - e)y(e^y - e) \end{aligned} \tag{3.23}$$

We will solve the problem of finding the variable parameter in 2D elliptic BVP (1.1). Here, we will formulate the regularized MOLS functional as a variational inequality problem of identifying $x^* \in A$ such that (3.1) holds.

3.3.1 Khobotov's Extragradient Method

In this section, we will state Marcotte's variations to reduce step length. The first one is known as the first modified version of Marcotte that keeps the α stay bounded away from zero:

$$\alpha = \min\left\{\frac{\alpha}{2}, \frac{1}{\sqrt{2}}\frac{||x^k - \bar{x}^k||}{||J'(x^k) - J'(\bar{x}^k)||}\right\} \tag{3.24}$$

The second modified version of Marcotte that is stated in [32] changes the initial α as following:

$$\alpha = \alpha_{k-1} + \left(\beta\frac{||x^{k-1} - \bar{x}^{k-1}||}{||J'(x^{k-1}) - J'(\bar{x}^{k-1})||} - \alpha_{k-1}\right)\gamma \tag{3.25}$$

where $\gamma \in (0,1)$ and $\beta \in (0,1)$. This rule increases the α if α_{k-1} is smaller than optimal, and the reduction rule for α is the following:

$$\alpha = \max\left\{\hat{\alpha}, \min\left\{\varepsilon\alpha, \beta\frac{||x^k - \bar{x}^k||}{||J'(x^k) - J'(\bar{x}^k)||}\right\}\right\} \tag{3.26}$$

where $\varepsilon \in (0,1)$. Therefore, these rules enable an adaptive alteration of the initial value of α_k.

Here is the Khobotov algorithm with Marcotte's two different choices to reduce step length:

Step 1: Choose initial $x^0 \in A, \beta, \alpha$ and set $k = 1$

(a) Update $\alpha = \alpha_{k-1}$ (Khobotov's version for step length)

(b) Update $\alpha = \alpha_{k-1} + \left(\beta\frac{||x^k - \bar{x}^k||}{||J'(x^k) - J'(\bar{x}^k)||} - \alpha_{k-1}\right)\gamma$ (Marcotte's second modified version)

Step 2: Compute $J'(x^k)$

Step 3: First Projection: $\bar{x}^k = P_A(x^k - \alpha J'(x^k))$ and compute $J'(\bar{x}^k)$

if $J'(\bar{x}^k) = 0$ *then* $\bar{x}^k$ is a solution of the problem.

else if $\alpha > \beta\frac{||x^k - \bar{x}^k||}{||J'(x^k) - J'(\bar{x}^k)||}$ then choose a rule to reduce step length:

(a) $\alpha = \min\left\{\frac{\alpha}{2}, \frac{1}{\sqrt{2}}\frac{||x^k - \bar{x}^k||}{||J'(x^k) - J'(\bar{x}^k)||}\right\}$ (Marcotte's first modified version)

(b) $\alpha = \max\left\{\hat{\alpha}, \min\left\{\varepsilon\alpha, \beta\frac{||x^k - \bar{x}^k||}{||J'(x^k) - J'(\bar{x}^k)||}\right\}\right\}$ (Marcotte's second modified version)

and go to Step 3.

else $\alpha_k = \alpha$

Step 4: Second Projection $x^{k+1} = P_A(x^k - \alpha J'(\bar{x}^k))$

Step 5: **if** $||x^{k+1} - x^k|| < TOL$, STOP, **else** $k = k+1$ go to Step 2.

The Figure 3.2 shows the result of computing the variable parameter for given 2D elliptic boundary value problem (3.23).

3.3.2 Solodov-Tseng Method

Solodov and Tseng [28] propose a practical alternative to the extragradient method, such that

$$\bar{x}^k = P_A(x^k - \alpha_k J'(x^k)) \qquad x^{k+1} = x^k - \gamma M^{-1}(T_\alpha(x^k) - T_\alpha P_A(\bar{x}^k)) \tag{3.27}$$

where γ is a positive step size, M is the scaling matrix that must be symmetric and positive definite, and $T_\alpha = (I - \alpha J')$; here I is the identity matrix. The extragradient method is modified by strongly monotone T_α and M^{-1}. Moreover, it only requires one projection, and two function evaluations at each step.
More details about the method can be found in [28, 32]. The following is the algorithm for the Scaled Extragradient Method that we use in our experiments:

Step 1: Choose initial $x^0 \in A, \theta \in (0,2), \rho \in (0,1), \alpha_0 > 0$ and set $k = 1$, positive, symmetric matrix, M.

Step 2: $\bar{x}^0 = 0$ and $rx = e$

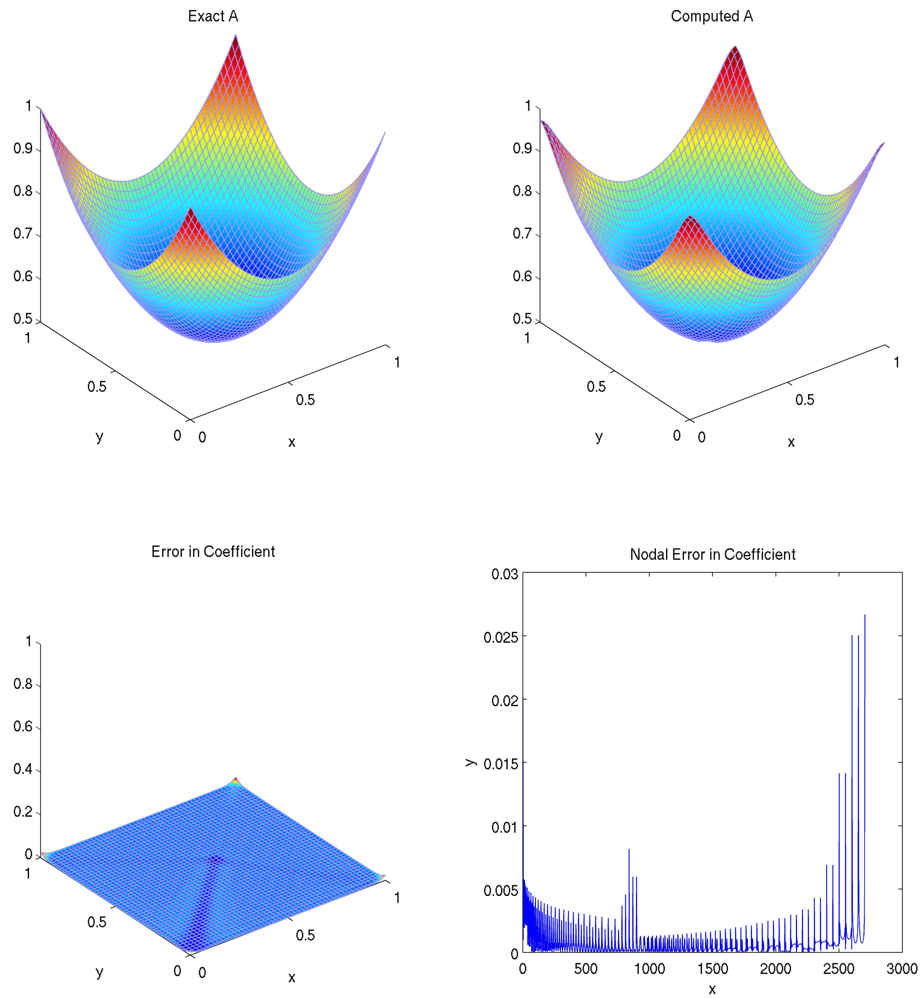


Figure 3.2: Solution by Second Modified Version of Marcotte

Step 3: **if** $||rx|| < TOL$ then STOP **else** $\alpha = \alpha_{k-1},\ flag = 0$

Step 4: **if** $J'(x^k) = 0$ then x^k is a solution of the problem.

Step 5: **while**

$$\alpha(x^k - \bar{x}^k)^T(J'(x^k) - J'(\bar{x}^k)) > (1-\rho)||x^k - \bar{x}^k||^2 \text{ or flag=0}$$

if flag$\neq 0$ then update $\alpha = \alpha_{k-1}\beta$ **endif**

update $\bar{x}^k = P_A(x^k - \alpha J'(x^k))$

compute $J'(\bar{x}^k)$

flag=flag+1;

endwhile

Step 6: Update $\alpha_k = \alpha$

Step 7: Compute $\gamma = \theta\rho||x^k - \bar{x}^k||^2/||M^{-1/2}(x^k - \bar{x}^k - \alpha_k J'(x^k) + \alpha_k J'(\bar{x}^k)||^2$

Step 8: Compute $x^{k+1} = x^k - \gamma M^{-1}(x^k - \bar{x}^k - \alpha_k J'(x^k) + \alpha_k J'(\bar{x}^k))$

Step 9: $rx = x^{k+1} - x^k$

$k = k+1$

go to Step 3.

endif

endif

The scaling matrix M can be taken as the identity matrix for simplification. The parameters θ, ρ, β has a vital role in the performance of the method. In our experiments, we chose various parameters to find the best approximation. The results can be found in the numerical experiments section.

Figure 3.3 shows the result of computing the variable parameter for given 2D elliptic boundary value problem (3.23).

3.3.3 Improved Goldstein's Method

Goldstein method was studied in [31] to solve variational inequality problems by updating the step length iteratively, that is:

$$x^{k+1} = P_A[x_k - \beta_k J'(x^k)] \tag{3.28}$$

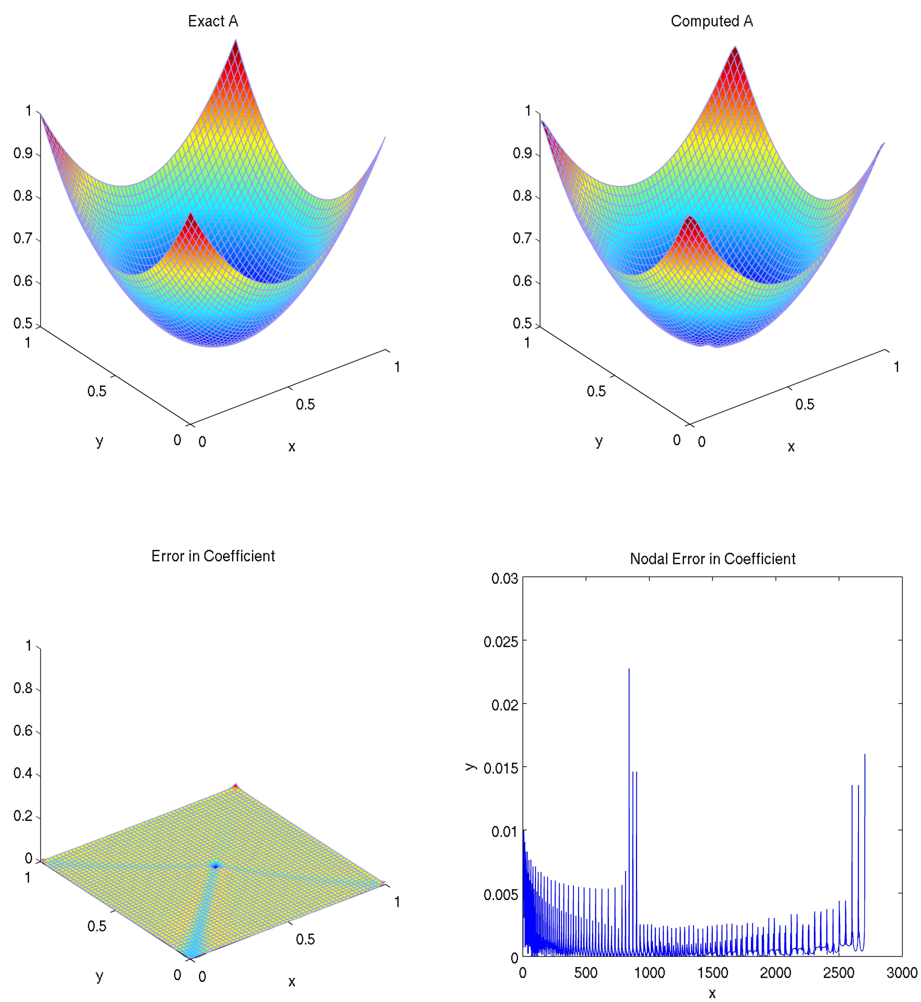


Figure 3.3: Solution by Solodov-Tseng Method

where β_k is positive scaling parameter. Moreover, [31] proposed the He-Goldstein Method which requires Lipschitz continuity and strong monotonicity of the objective function.

$$x^{k+1} = x^k - \frac{1}{\beta_k}\left\{J'(x^k) - P_A[J'(x^k) - \beta_k x^k]\right\}. \tag{3.29}$$

Furthermore, an improved version of the Goldstein's method is provided in [31] such that

$$\begin{aligned} x^{k+1} &= x_k - \gamma\alpha_k r(x^k,\beta_k) \\ r(x^k,\beta_k) &= \frac{1}{\beta_k}\left\{J'(x^k) - P_A[J'(x^k) - \beta_k x^k]\right\}. \end{aligned} \tag{3.30}$$

where $\alpha = 1 - \frac{1}{4\beta_k\tau}$ and $\gamma \in (0,2)$.

We are going to use (3.30) for our numerical experiments. The step algorithm is the following:

Step 1: Choose initial $x^0 \in A, \varepsilon > 0$, $\gamma \in (0,2), \beta_U > \beta_L > 1/(4\tau)$ where $\beta_0 \in [\beta_L, \beta_U]$ and set $k = 0$.

Step 2: Compute

$$r(x^k,\beta_k) = \frac{1}{\beta_k}\left\{J'x^k) - P_A[J'(x^k) - \beta_k x^k]\right\}.$$

If $||r(x^k,\beta_k) \leq \varepsilon$ then x_k is a solution to the problem.

Step 3: Compute the next iteration $x^{k+1} = x_k - \gamma\,\alpha_k\, r(x^k,\beta_k)$

Step 4: Update β_k

if $\frac{||J'(x^{k+1})-J'(x^k)||}{\beta_k||x^{k+1}-x^k||} < \frac{1}{2}$ set $\beta_{k+1} = \max\{\beta_L, \frac{1}{2}\}\beta_k$

else if $\frac{||J'(x^{k+1})-J'(x^k)||}{\beta_k||x^{k+1}-x^k||} > \frac{3}{2}$ set $\beta_{k+1} = \min\{\beta_U, \frac{6}{5}\}\beta_k$

Step 5: Set $k = k+1$ and go to step 2.

Figure 3.4 shows the result of computing the variable parameter for given 2D elliptic boundary value problem (3.23).

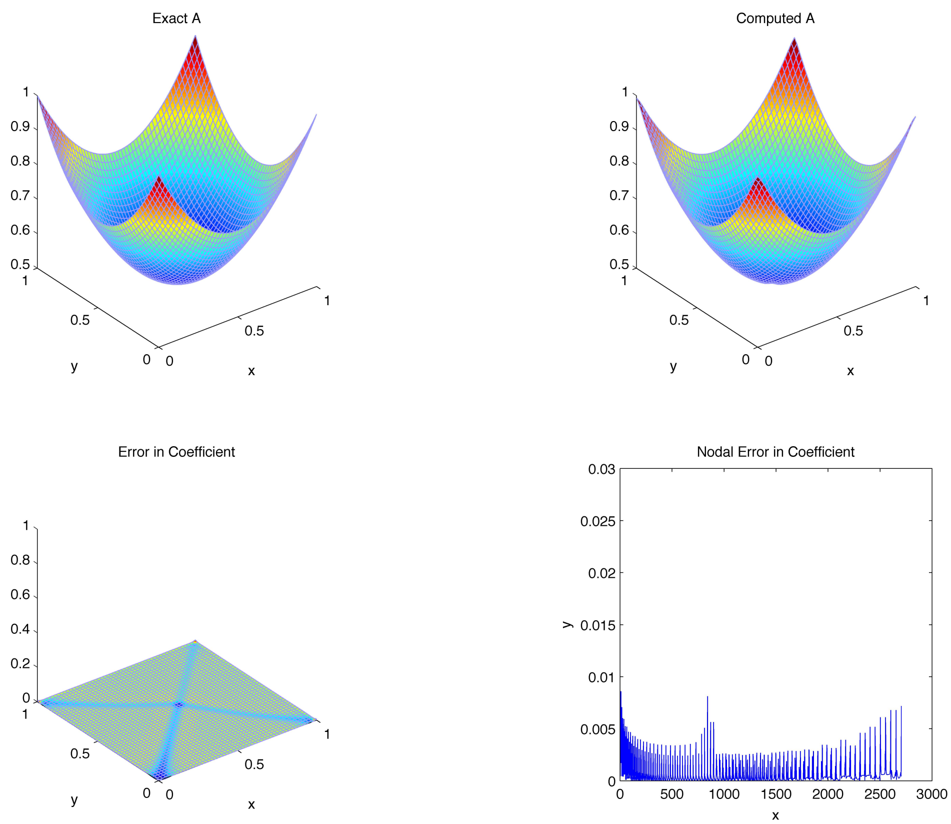


Figure 3.4: Solution by Improved Goldstein's Method

3.3.4 Hyperplane Method

Iusem [16] introduced a geometric interpretation of the extragradient methods. The idea behind this algorithm is to use a hyperplane to separate the solutions of the variational problem to the one side. The projection method is described in (3.22). Here, if we fix $\alpha_k = \lambda_k = \alpha$, then we obtain a similar iteration to Korpelevich's extragradient method [23].

We will need to choose $\varepsilon \in (0,1)$, $\hat{\alpha}$,and $\tilde{\alpha}$ such that $0 < \hat{\alpha} < \tilde{\alpha}$ at initial step. The step algorithm is the following [16]:

Step 1: Choose initial $x^0 \in A, \varepsilon \in (0,1), ,0 < \hat{\alpha} < \tilde{\alpha}$

Step 2: **if** $P_A(x^k - \alpha_k J'(x^k)) = x^k$ then stop
else select the step length

$$\textbf{if } ||J'(\bar{x}^k) - J'(x^k)|| \le \frac{||\bar{x}^k - x^k||}{2\tilde{\alpha}_k{}^2 ||J'(x^k)||} \textbf{ then } \bar{x}^k = \tilde{x}^k$$

$$\textbf{else } find\ \ \alpha_k \in (0, \tilde{\alpha}_k)\ \ such\ that$$

$$\varepsilon \frac{||\bar{x}^k - x^k||}{2\tilde{\alpha}_k{}^2 ||J'(x^k)||} \le ||J'(P_A(x^k - \alpha_k J'(x^k))) - J'(x^k)|| \le \frac{||\bar{x}^k - x^k||}{2\tilde{\alpha}_k{}^2 ||J'(x^k)||}$$

endif

$$\textbf{if} J'(\bar{x}^k) = 0 \ \ \textbf{then } \bar{x}^k \in A^*\ \ stop$$

$$\textbf{else } compute\ \ the\ new\ iterate$$

$$x^{k+1} = P_A\left(x^k - \frac{\langle J'(\bar{x}^k, x^k - \bar{x}^k\rangle}{||J'(\bar{x}^k)||^2} J'(\bar{x}^k)\right)$$

Step 3: Set $k = k+1$ and go to step 2.

Figure 3.5 shows the result of computing the variable parameter for given 2D elliptic boundary value problem (3.23).

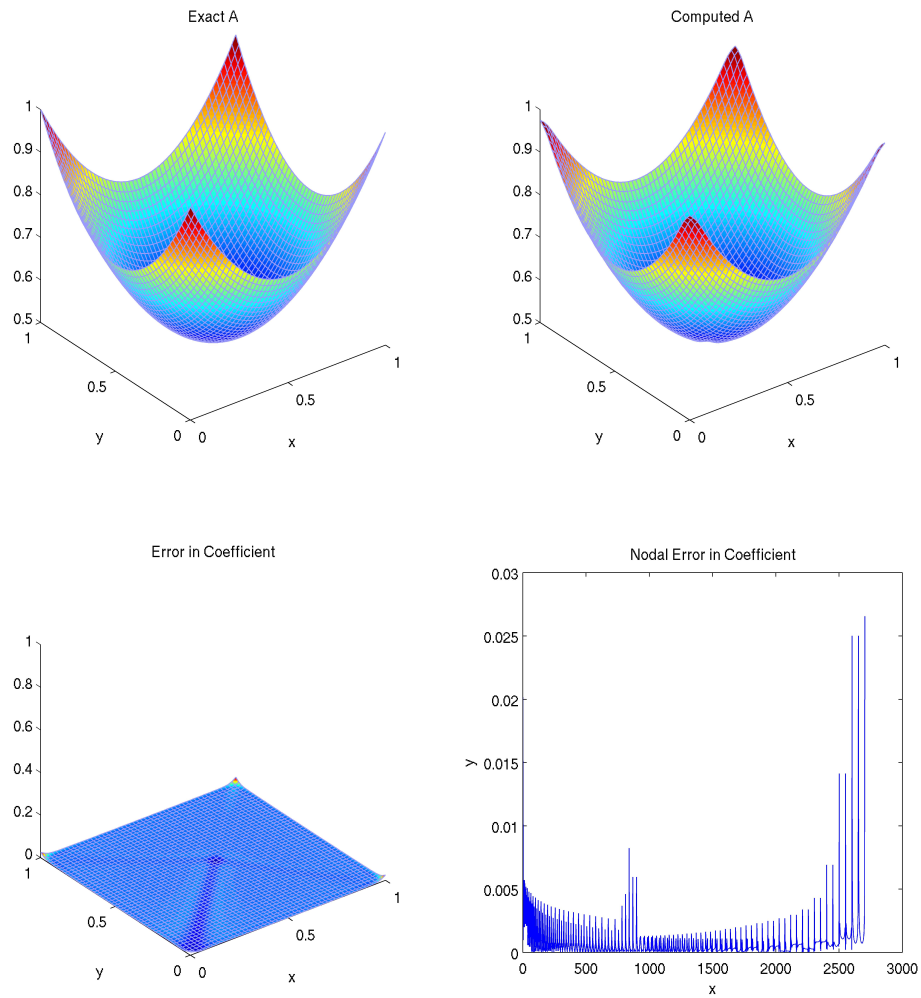


Figure 3.5: Solution by Hyperplane Method

Chapter 4

Performance Analysis

In this section, we will give a brief explanation for the performance analysis of each method. We will emphasize the importance of choosing step length and how it affects the efficiency of the extragradient method. Moreover, each algorithm has constant parameters that we have to declare at the initial step. These parameters play a significant role for the implementation of the extragradient methods. We will use MOLS functional for our test problem unless otherwise stated.

In the following Figure 4.1 and Figure 4.2, we show how the step length α_k alters depending on the extragradient methods, Solodov-Tseng and Improved Goldstein's method, respectively.

As we described in the previous chapter, Khobotov [18] uses a constant step length for the extragradient method. However, the value of the parameter affects how fast the algorithm converges. For example, if we set $\alpha = 0.1$, the gradient norm decreases from 0.0048 to 0.0047 in nearly 400 iterations and in 10000 iterations, the gradient norm becomes 0.0030. On the other hand, for $\alpha = 100$, the gradient norm decreases from 0.0048 to $2.9221e-5$ in nearly 400 iterations. However, the method does not converge when the step length is too large.

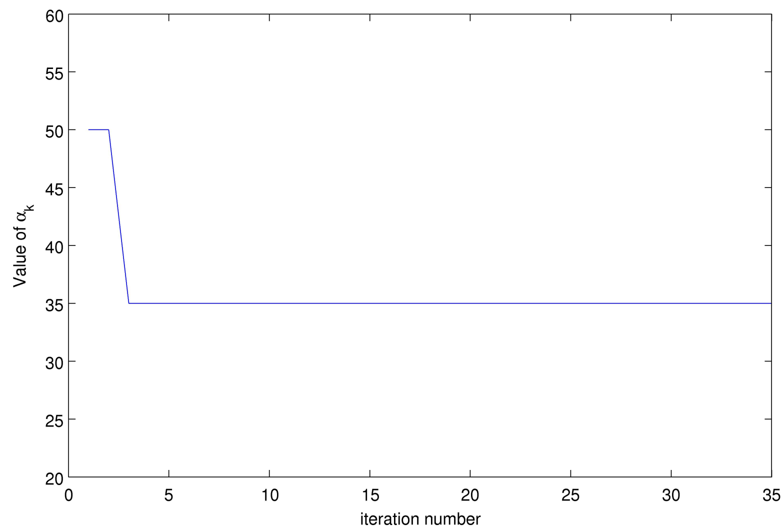


Figure 4.1: Reduction Rule for Solodov-Tseng Method

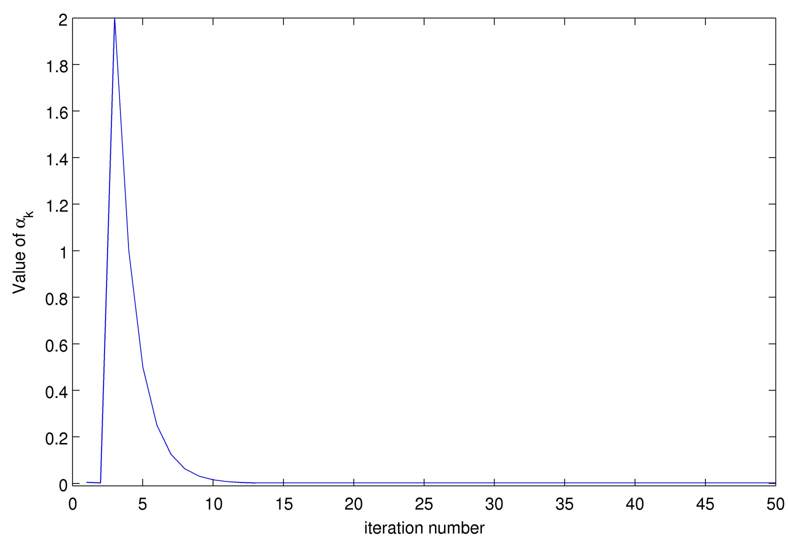


Figure 4.2: Reduction Rule for Improved Improved Goldstein's Method

The first modified version of the Marcotte rule cannot increase the value of the step length whereas the second modified version of Marcotte can increase the α_k if it is smaller than optimum. Moreover, it is also easy to observe that

α	0.1	100
L2-Error	0.0653	1.1890e-05
H1 error	1.1602	0.0779
2-norm	7.9882	0.1673

Table 4.1: α performance by Khobotov

Solodov and Tseng applies a reduction rule for step length by multiplying α_k by a constant β while the condition that is stated in the algorithm is satisfied.

We solved the same test problem by using two different objective functionals which are OLS and MOLS. We computed the minimum eigenvalues of the hessian matrix at each iteration for both functionals. Figure 4.3 and Figure 4.4 illustrates the results for the first 200 iterations.

It is useful to state the minimum eigenvalue using MOLS functional is positive at each iteration which clarifies that MOLS is convex. On the other hand, we obtain negative eigenvalues by using OLS function. We implement MOLS functional on our test example since it is known that the MOLS is convex [12] .

Table 4.2 gives the error analysis for various initial step length implementing the second modified version of Marcotte method with MOLS functional.

α	Iteration	L2-Error	H1 error	2-norm
0.001	1281	1.5378e-06	0.0070	0.0730
10	1281	1.5375e-06	0.0070	0.0730
100	1269	1.531e-06	0.0069	0.0729
100	1137	1.9938e-06	0.0084	0.0832

Table 4.2: α performance by MOLS

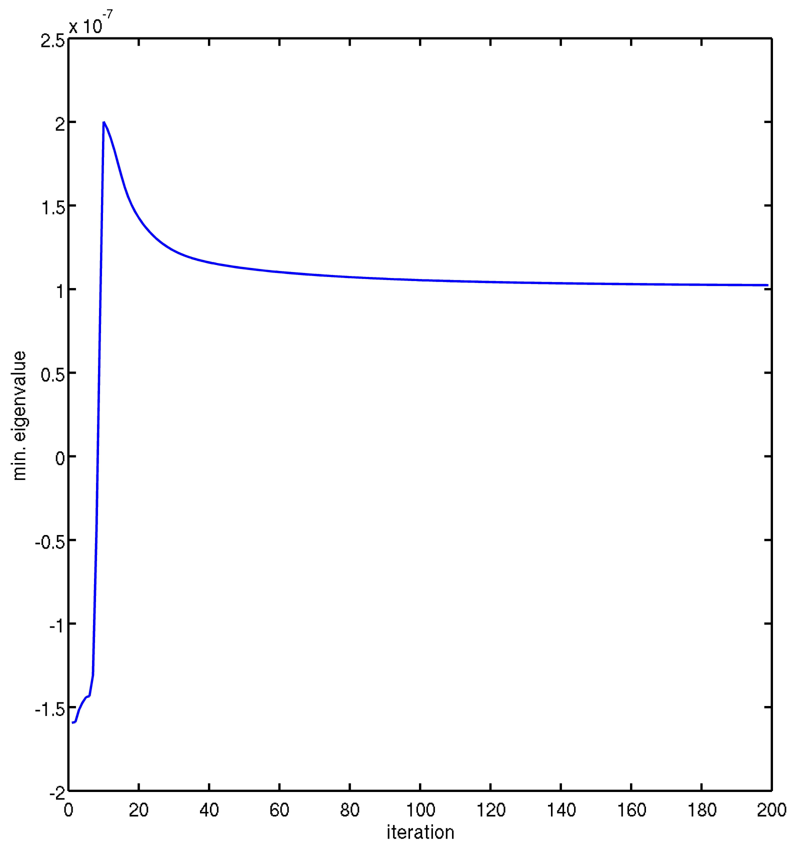


Figure 4.3: Performance Analysis of OLS

Table 4.3 gives the error analysis for various initial step length implementing the second modified version of Marcotte method with OLS functional.

We already mentioned the importance of choosing regularization parameter. Table 4.4 gives L_2 error analysis for different regularization parameters by using Marcotte (SMV).

To demonstrate, we can refer to Figure 4.5 that shows how the error gets smaller when we get closer to the suitable regularization parameter. We can also illustrate the definition of under-regularized and over-regularized solution by referring to (a) and (c) of Figure 4.5, respectively. Here, we include the com-

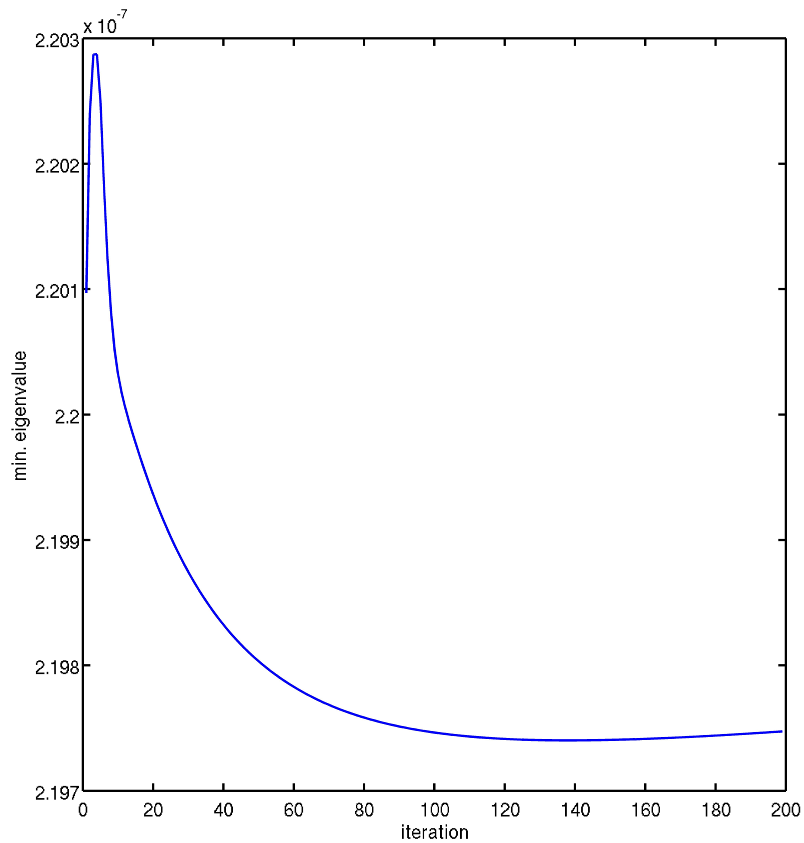


Figure 4.4: Performance Analysis of MOLS

α	Iteration	L2-Error	H1 error	2-norm
0.001	771	4.2847e-04	0.1965	0.7200
10	771	4.2848e-04	0.1965	0.7200
100	772	4.2696e-04	0.1961	0.7186
100	558	3.6831e-04	0.1926	0.6486

Table 4.3: α performance by OLS

ε	L2-Error	Iteration
2e-03	0.0011	1260
2e-05	1.7932e-05	504
2e-06	1.1543e-06	1143
2e-08	4.8248e-05	2306

Table 4.4: Regularization parameter, ε, performance

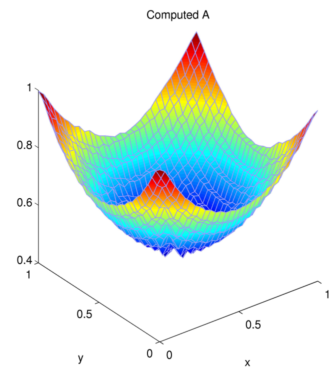
puted solution for different ε values. When $\varepsilon = 2E-8$, the solution is under-regularized which states that we need to implement a bigger value for regularization parameter. On the other hand, $\varepsilon = 2E-5$ gives an over-regularized solution which clarifies that a smaller regularization parameter can give a better result. Hence, $\varepsilon = 2E-6$ gives the best approximation for Marcotte (SMV) method.

We implement the extragradient methods for our test problem from the previous chapter. Here, we give the suitable regularization parameter for each method and outputs for $L_2 - error,\ and\ H_1 - error$.

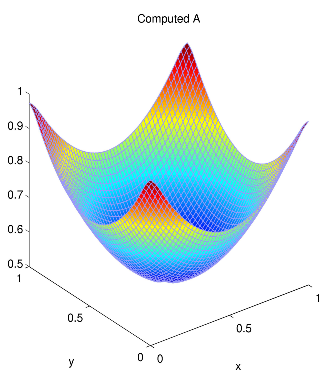
Extragradient Method	Iteration	ε	L2-Error	H1 error
Marcotte-SMV	1143	2e-06	1.1543e-06	0.0066
Solodov-Tseng	2150	8e-07	1.9382e-06	0.0077
Improved Goldstein's	61907	8e-07	1.2149e-06	0.0031
Hyperplane	22511	2e-06	1.1399e-06	0.0066

Table 4.5: Performance analysis for the test problem

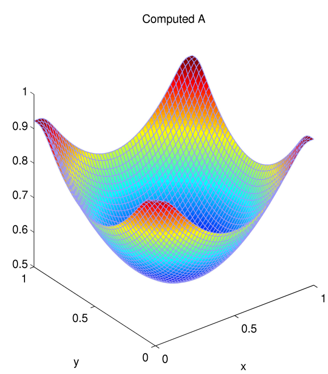
Computed solutions and errors by the extragradient methods can be found in the previous chapter. Figure 4.6 demonstrates the computed solution by each extragradient method. Here, we implemented the extragradient methods for the



(a) $\varepsilon = 2E - 8$



(b) $\varepsilon = 2E - 6$



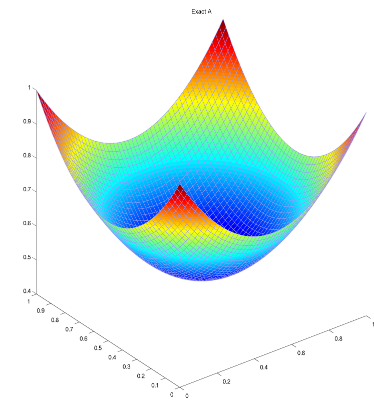
(c) $\varepsilon = 2E - 5$

Figure 4.5: Solution by Different Regularization Parameters

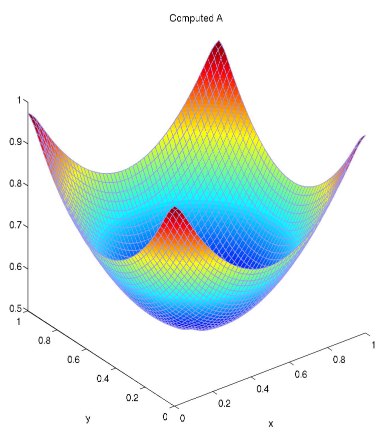
test problem with n=50.

We can conclude that we get the best approximation by using Marcotte methods since these methods have few parameters that we need to declare at initial step. However, to choose the right step length at initial step still has a significant role for the performance of the method. Solodov and Tseng introduces a scaled matrix to accelerate the convergence. Moreover, Improved Goldstein's method is also very effective method. This method is also sensitive with the parameters since choosing poor initial parameters may lead to over or under regularized solutions.

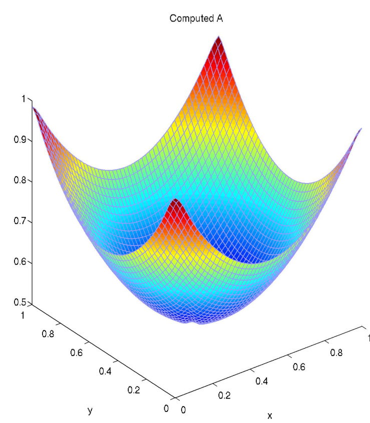
Overall, changing only one parameter affects the entire algorithm and the accuracy of the computed solution. Hence, we get better results with Marcotte methods since it requires few parameters to be declared at initial step. That is the reason why we mainly implemented Marcotte methods for performance analysis.



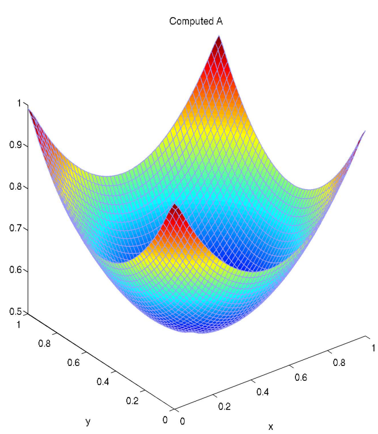
(a) Exact



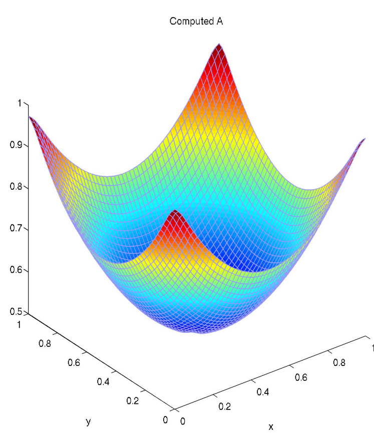
(b) Marcotte



(c) Solodov-Tseng



(d) Improved Goldstein's



(e) Hyperplane

Figure 4.6: Computed A for Various Extragradient Methods

Chapter 5

Background Material

Vector spaces and inner product play a significant role in understanding the finite element methods. A *vector space* is a set of vectors V on which two algebraic operations are defined:
Vector addition: $u, v \in V$ then $u + v \in V$. It is clear to see that the vector addition is commutative and associative.
Scalar multiplication: $u \in V$ and $\alpha \in \mathrm{R}$ then $\alpha u \in V$.

It is important to note that functions can be written as vectors, thus they satisfy the fundamental properties of vectors.

Definition 5.0.4 (Vector Space)**.** *Let V be a vector space. $(\cdot, \cdot)$ be an inner product on V obtaining a real number by taking two vectors from V that holds the following properties:*

- $(u, v) = (v, u)$ *for all* $u, v \in V$
- $(\alpha u + \beta v, w) = \alpha(u, w) + \beta(v, w)$ *for all* $u, v, w \in V$ *and all* $\alpha, \beta \in \mathrm{R}$
- $(u, u) \geq 0$ *for all* $u \in V$*, and* $(u, u) = 0$ *if and only if* $u = 0$.

It can be concluded that $(w, \alpha u + \beta v) = \alpha(w, u) + \beta(w, v)$ *also holds for all* $u, v, w \in V$ *and all* $\alpha, \beta \in \mathrm{R}$.

Definition 5.0.5 (Norm Space)**.** *Let V be a vector space $||\cdot||$ be a norm on V with a real-valued function defined on V that holds the following properties:*

- $||u|| \geq 0$ *for all* $u \in V$*, and* $||u|| = 0$ *if and only if* $u = 0$*;*
- $||\alpha u|| = |\alpha|||u||$ *for all* $u \in V$ *and all* $\alpha \in \mathrm{R}$
- $||u+v|| \leq ||u|| + ||v||$ *for all* $u, v \in V$*.*

Therefore, we can define a norm on V as $||u|| = \sqrt{(u,u)}$ where (u,u) is an inner product on V.

Definition 5.0.6 (Banach Space)**.** *Let $\{X_n\}$ be a Cauchy sequence in a normed space X. If every Cauchy Sequence converges to a limit in X, then X is called complete and every complete normed space is called a Banach space.*

Theorem 5.0.7. *Let V be an inner product space, then for all $u, v \in V$ the following properties hold:*

1. $|(x,y)| \leq ||x||\,||y||$
2. $||x+y|| \leq ||x|| + ||y||$
3. $||x+y||^2 + ||x-y||^2 = 2(||x||^2 + ||y||^2)$

Definition 5.0.8 (Linear Operators)**.** *Let X,Y be real vector spaces and define a map $A : X \to Y$ where*

$$A(\alpha u + \beta v) = \alpha A u + \beta A v$$

holds for $\forall u, v \in D(A)$ and $\alpha, \beta \in \mathrm{R}$, then A is called a linear operator. A is called bounded if $||Au|| \leq m||u||$ $\forall u \in X$ for a constant $m > 0$.

Let $A(u,v)$ be a symmetric bilinear form than the following properties will be hold:

1. $a(u,v) = a(v,u)$ *for all* $u, v \in V$*.*

2. $a(\alpha u+\beta v,w)=\alpha a(u,w)+\beta a(v,w)$ *for all* $u,v,w \in V$ *and all* $\alpha,\beta \in \mathrm{R}$

3. $a(u,u)\geq 0$, *and* $u=0$ *then* $a(u,u)=0$
   Moreover, in some cases, the bilinear form $a(\cdot,\cdot)$ *can also hold the following properties:*

4. *There exists* $\alpha > 0$ *such that* $a(u,u)\geq \alpha||u||^2$ *for all* $u\in V$. *(V-elliptic)*

5. *There exists* $\beta > 0$ *such that* $a(u,v)\leq \beta||u||\,||v||$ *for all* $u,v\in V$.*(Bounded)*

Definition 5.0.9 (Dual Space). X^* *is called dual space consisting of all linear functionals on a normed space X that is defined by*

$$||f|| = \sup_{x\in X, x\neq 0} \frac{|f(x)|}{||x||}$$

Definition 5.0.10 (Hilbert Spaces). *Let H be a Hilbert space, the it satisfies the followings*

1. *H is a vector space.*

2. *H is an inner product space that is* $(\cdot,\cdot): H\times H \to K$ *(K is a scalar field) such that*

   i $(u,v)=(\overline{v,u})$

   ii $(\alpha x+\beta y,z)=\alpha(x,z)+\beta(y,z)$

   iii $(x,\gamma y+\delta z)=\bar{\gamma}(x,y)+\bar{\delta}(x,z)$

   iv $(u,u)\geq 0$ *with equality holding if and only* $u=0$

3. *H is a normed space such that* $||\cdot||=\sqrt{(\cdot,\cdot)}$

4. *H is complete.*

Hence, a vector space H with an inner product $(\cdot,\cdot)$ *that is complete with respect to norm* $||\cdot||=\sqrt{(\cdot,\cdot)}$ *is called a Hilbert space.*

Now, we can state two well-known inequalities, that are

1. Cauchy-Schwarz Inequality $|(x,y)| \leq ||x||\,||y||$

2. Minkowski Inequality $||x+y|| \leq ||x|| + ||y||$

Theorem 5.0.11 (Riesz Theorem)**.** *Let H be a Hilbert space, then every $f \in H^*$ can be represented by inner product, that is*

$$f(x) = (x,y)$$

where $||f|| = ||y||$.

Theorem 5.0.12 (The Lax-Milgram theorem)**.** *Let V be Hilbert space and $a(u,v)$ be V-elliptic, bounded bilinear form such that*

$$|a(u,v)| \leq \beta ||u||\,||v|| \ for\,all\, u,v \in V, \beta > 0$$

$$||a(v,v)|| \geq \alpha ||v||^2 \ v \in V, \alpha > 0$$

Let $f \in V^*$, then there exist a unique solution $u \in V$ to the variational problem such that

$$a(u,v) = f(v) \ for\,all\, v \in V$$

and the solution u continuously depends on f such that

$$||u||_V \leq \frac{1}{\alpha} ||f||_V^*$$

Definition 5.0.13. *Define an operator $A : X \to X^*$ on the real Banach space X.*
Monotone: $(Au - Av, u - v) \geq 0 \ for\ all\ u,v \in V$

Strongly Monotone: $(Au - Av, u - v) \geq c||u-v||_x^p \ for\,all\, u,v \in X, where\ c > 0\ and\ p > 1$

Coercive: $\lim_{||u|| \to \infty} \frac{(Au,u)}{||u||} = +\infty$ *A strongly monotone operator is coercive.*

Convexity: A functional $f : A \subset X \rightarrow \mathrm{R}$ *is a convex set in the real normed space* X *that is:* $f((1-t)u+tv) \leq (1-t)f(u)+tf(v)$ $for\ all$ $t \in [0,1]$ $u,v \in A$

Lemma 5.0.14. *Let* H *be a real Hilbert space, and* A *be a nonempty, closed , convex subset of* H*. For each* $x \in H$ *there is a unique* $y \in A$ *such that*

$$||x-y|| = \inf_{z \in A} ||x-z|| \tag{5.1}$$

where y *is called the projection of* x *on* A *such that*

$$y = P_A x$$

Definition 5.0.15. *Let* M *be a metric space and* $F : M \rightarrow M$ *is called contraction mapping if*

$$d(F(x),F(y)) \leq \hat{k}\, d(x,y), \ \ x,y \in M$$

for some $0 < \hat{k} < 1$*. If* $\hat{k} = 1$ *then* M *is called nonexpansive.*

Theorem 5.0.16. *The projection of* x *on* A*,* $y = P_A x$ *where* A *is a closed convex set of Hilbert space, if and only if:*

$$y \in A : \langle y, z-y \rangle \geq \langle x, z-y \rangle \text{ for all } \ z \in A$$

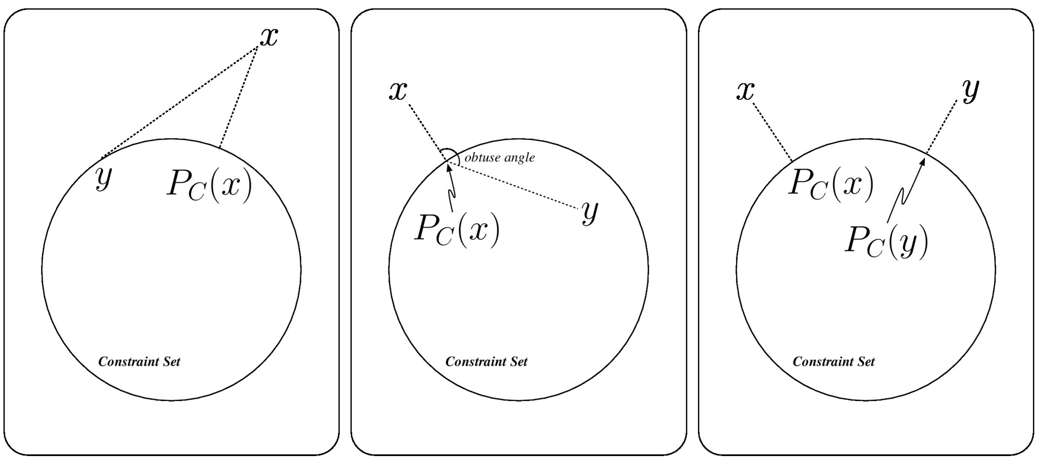
Lemma 5.0.17. *The projection operator satisfies the following properties [19]:*

(a) $||P_A x - P_A y|| \leq ||x-y||$ *for all* $x,y \in H$

(b) $\langle x - P_A x, P_A x - y \rangle \geq 0$ *for all* $x \in H,\ y \in A$

(c) $||x-y||^2 \geq ||x-P_A x||^2 + ||y-P_A x||^2$ *for all* $x \in H,\ y \in A$

The geometric representation is shown for the above properties in Figure 5.1.



(a) (b) (c)

Figure 5.1: The Geometric Representation of the Projection Properties

Bibliography

[1] R. Acar, Identification of the coefficient in elliptic equations, *SIAM J. Control Optim.*, 31, (1993), 1221–1244.

[2] A. S. Antipin, B. A. Budak, F. P. Vasilév, A first-order continuous extragradient method with a variable metric for solving equilibrium programming problems. (Russian) Vestnik Moskov. Univ. Ser. XV Vychisl. Mat. Kibernet. 2003, no. 1, 37–41, 56; translation in Moscow Univ. Comput. Math. Cybernet. 2003, no. 1, 43–47

[3] A. Bnouhachem, M. Aslam Noor, Z. Hao, Some new extragradient iterative methods for variational inequalities. Nonlinear Anal. 70 (2009), no. 3, 1321–1329.

[4] A. Bnouhachem, Abdellah; X. Fu, M. H. Xu, and S. Zhaohan,Modified extragradient methods for solving variational inequalities. Comput. Math. Appl. 57 (2009), 230–239.

[5] A. Bnouhachem, Abdellah; X. Fu, M. H. Xu, and S. Zhaohan, New extragradient-type methods for solving variational inequalities. Appl. Math. Comput. 216 (2010), 2430–2440.

[6] Y. Censor, A. Gibali, S. Reich, The subgradient extragradient method for solving variational inequalities in Hilbert space. J. Optim. Theory Appl. 148 (2011), no. 2, 318-335.

[7] Y. Censor, A. Gibali, S. Reich, Strong convergence of subgradient extragradient methods for the variational inequality problem in Hilbert space. Optim. Methods Softw. 26 (2011), 827–845.

[8] Y. Censor, A. Gibali, S. Reich, Extensions of Korpelevich's extragradient method for the variational inequality problem in Euclidean space. Optimization 61 (2012), 1119–1132.

[9] M. S. Gockenbach, Partial Differential Equations Analytical and Numerical Methods. SIAM, (2002).

[10] M. S. Gockenbach, Understanding and Implementing the Finite Element Method. SIAM, (2006).

[11] M. S. Gockenbach, A. A. Khan, Identification of Lamé parameters in linear elasticity: a fixed point approach, J. Ind. Manag. Optim. 1 (2005)487–497.

[12] M. S. Gockenbach, A. A. Khan, An abstract framework for elliptic inverse problems. Part 1: an output least-squares approach, Math. Mech. Solids, 12 (2007) 259–276.

[13] M. S. Gockenbach, B. Jadamba, A. A. Khan, Numerical estimation of discontinuous coefficients by the method of equation error, Int. J. Math. Comput. Sci., 1 (2006) 343–359.

[14] M. S. Gockenbach, B. Jadamba, A. A. Khan, Equation error approach for elliptic inverse problems with an application to the identification of Lamé parameters, Inverse Problems in Science and Engineering, 16 (2008) 349–367.

[15] M. S. Gockenbach, B. Jadamba, A. A. Khan, A comparative numerical study of optimisation approaches for elliptic inverse problems, JMI Inter. J. Math. Sci.,1,(2010), 35–56.

[16] A.N. Iusem, An iterative algorithm for the variational inequality problem, Computational and Applied Mathematics ,Vol. 13, N.2, (1994), pp. 103-114.

[17] A. N. Iusem and B.F. Svaiter, A variant of Korpelevich's method for variational inequalities with a new search strategy,*Optimization,* Vol. 42, 1997, 309–321.

[18] Khobotov, E. N. A modification of the extragradient method for solving variational inequalities and some optimization problems. (Russian) Zh. Vychisl. Mat. i Mat. Fiz. 27 (1987), no. 10, 14620–1473.

[19] D. Kinderleher and G. Stampacchia, An Introduction to Variational Inequalities and Their Applications. *SIAM Classics In Applied Mathematics*, 31, (2000).

[20] Kluge, R. (1979): *Nichtlineare Variationsungleichungen und Extremalaufgaben,* Mathematische Monographien, 12. VEB Deutscher Verlag der Wissenschaften, Berlin.

[21] Kluge, R. (1983): Zur "Koeffizienten"bestimmung in linearen Operator- und Evolutionsgleichungen, *Math. Nachr.*, 112 , 153–175.

[22] I. Knowles, Parameter identification for elliptic problems, J. Comp. Appl. Math., 131 (2001) 175–194.

[23] G. M. Korpelevic, An extragradient method for finding saddle points and for other problems. (Russian) Ekonom. i Mat. Metody 12 (1976), no. 4, 747–756.

[24] P. Marcotte, Application of Khobotov's algorithm to variational inequalities and network equilibrium problems, Inform. Systems Oper. Res., 29 (1991), 258–270.

[25] R. D.C. Monteiro, B. F. Svaiter, Iteration-complexity of a Newton proximal extragradient method for monotone variational inequalities and inclusion problems. SIAM J. Optim. 22 (2012), no. 3, 914-935.

[26] Oleksyn, J. (2010, June). Extragradient methods for elliptic inverse problems and image denoising (Masters thesis).

[27] L. D. Popov, L. D. On schemes for the formation of a master sequence in a regularized extragradient method for solving variational inequalities. (Russian) Izv. Vyssh. Uchebn. Zaved. Mat. 2004, no. 1, 70–79; translation in Russian Math. (Iz. VUZ) 48 (2004), no. 1, 6776

[28] M.V. Solodov, P. Tseng Modifield projection type methods for monotone variational inequalities SIAM Journal Control Optimization, Vol.34, N.5, (1996), pp. 1814–1830.

[29] M. V. Solodov, B. F. Svaiter, A new projection method for variational inequality problems, SIAM J. Control Optim. 37 (1999), 765–776.

[30] B.S. He, A Goldsteins type projection method for a class of variant variational inequalities, J. Comput. Math. 17 (4) (1999) 425-434.

[31] M.Li, L.-Z. Liao, X.M. Yuan, Some Goldstein's type methods for cocoercive variant variational inequalities. Appl. Numer. Anal. 61 (2011), 216–228.

[32] F. Tinti, F. Numerical solution for pseudomonotone variational inequality problems by extragradient methods. Variational analysis and applications, 1101–1128, Nonconvex Optim. Appl., 79, Springer, New York, 2005.

[33] A. J. Zaslavski, The extragradient method for solving variational inequalities in the presence of computational errors, J. Optim. Theory Appl., 153 (2012), 602–618.

[34] A. J. Zaslavski, The extragradient method for convex optimization in the presence of computational errors. Numer. Funct. Anal. Optim. 33 (2012), 1399–1412.

[35] A. J. Zaslavski, The extragradient method for finding a common solution of a finite family of variational inequalities and a finite family of fixed point problems in the presence of computational errors. J. Math. Anal. Appl., 400 (2013), 651–663.